



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/761,451	01/20/2004	Dong Yu	M61.12-0582	3046
27366 7590 12/05/2008 WESTMAN CHAMPLIN (MICROSOFT CORPORATION) SUITE 1400 900 SECOND AVENUE SOUTH MINNEAPOLIS, MN 55402-3244				
EXAMINER SHAH, PARAS D				
ART UNIT		PAPER NUMBER		
2626				
MAIL DATE		DELIVERY MODE		
12/05/2008		PAPER		

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

### Office Action Summary

**Application No.**

10/761,451

**Applicant(s)**

YU ET AL.

**Examiner**

PARAS SHAH

**Art Unit**

2626

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 04 September 2008.  
2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.  
3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1, 3, 4, 6, 7, 14-19, 21-23 is/are pending in the application.  
4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.  
5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.  
6) ☒ Claim(s) 1, 3, 4, 6, 7, 14-19, 21-23 is/are rejected.  
7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.  
8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.  
10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).  
11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
a) ☐ All b) ☐ Some \* c) ☐ None of:  
1. ☐ Certified copies of the priority documents have been received.  
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)  
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)  
3) ☐ Information Disclosure Statement(s) (PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_  
4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date \_\_\_\_\_  
5) ☐ Notice of Informal Patent Application  
6) ☐ Other: \_\_\_\_\_

### **DETAILED ACTION**

1. This communication is in response to the Amendments and Arguments filed on 09/04/2008. Claims 1, 3, 4, 6, 7, 14-19, and 21-23 remain pending and have been examined. The Applicants' amendment and remarks have been carefully considered, but they do not place the claims in condition for allowance. Accordingly, this action has been made FINAL.
2. All previous objections and rejections directed to the Applicant's disclosure and claims not discussed in this Office Action have been withdrawn by the Examiner.

### ***Response to Amendments and Arguments***

3. Applicant's arguments (pages 5-7) filed on 09/04/2008 with regard to claims 1 and 7 have been fully considered but they are moot in view of new grounds for rejection.

### ***Claim Objections***

4. Claim 7 is objected to because of the following informalities: "the pronunciation" in lines 12 should be "a pronunciation". Appropriate correction is required.
5. Claims 14-19 and 31-23 are objected to as being dependent upon an objected to base claim.

### ***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 1, 3, 4, 6, 7, and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nassiff *et al.* in view of Hon *et al.* (US 5,852,801).

As to claim 1, Nassiff *et al.* teaches

a computer-implemented speech recognition system comprising:

a microphone to receive user speech (see col. 4, lines 16-18);

a speech recognition engine coupled to the microphone (see col. 4, lines 16-17) (e.g. The speech recognition engine receives input from the microphone so it is implied that the two are coupled.), and being adapted to recognize the user speech (see col. 4, lines 15-19) and provide a textual output on a user interface (see col. 2, lines 19-20 and col. 5, lines 32-38); and

wherein the recognition engine is adapted to determine if the user's pronunciation caused the error, and selectively modify a probability associated with an existing pronunciation (see col. 7, lines 55-66) (e.g. The use of a statistical quantity with the updating of a language model implies that a probability value is associated with a word when comparisons are made (see col. 6, lines 28-31)).

However, Nassiff does not specifically teach the selectively increase a probability associated with word.

Hon does teach selectively increase a probability associated with word (see col. 2, lines 30-36, where the language model is adapted to increase the chance the same word is recognized by increasing the unigram probability.)

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the correction of dictating speech of Nassiff *et al.* with the inclusion of increasing a probability as taught by Hon (801). The motivation to have combined the references involves the reduction of errors when spoken words are not found in the lexicon of the recognition engine so as to adapt to unrecognized words in a speech recognition system (see Hon (801) col. 1, lines 33-36 and lines 54-56).

As to claims 3 and 21, Nassiff in view of Hon teaches all of the limitations as in claim 1

Furthermore, Nassiff *et al.* teaches the use of a user lexicon (see col. 6, line 25 and col.6, line 28)) (e.g. the alternative word list). .

Hon *et al.* (801) does teach the use of a lexicon, which is updated for new words (see col. 9, lines 36-40), where words are added when determining if the words exist in the user lexicon (see col. 7, lines 66-67 and col. 8, lines 1-3) (e.g. The determination is made of whether the word is in the lexicon if it is unrecognized). (e.g. Since the language model is updated the temporary storing of words in Nassiff based on presence or absence in the user lexicon would be obvious to one of skilled in the art. Further, it was stated that the word "two

much" and "too much" is added to the lexicon, where the words two and too are a word pair. Hence, Nassiff teaches a similar word pair being step and steep. The misrecognition of step to be steep would be a word pair when added to the list of words as taught by Hon.)

As to claim 4, Nassiff in view of Hon teaches all of the limitations as in claim 1.

Furthermore, Nassiff *et al.* teaches wherein the recognition engine is adapted to determine if the user's pronunciation caused the error and selectively learn the new pronunciation (see col. 6, lines 45-50 and lines 57-58) (e.g. The determination is made as to whether a misrecognition error has occurred, if so the language model is updated.).

As to claim 6, Nassiff in view of Hon teaches all of the limitations as in claim 1.

Furthermore, Nassiff *et al.* teaches the updating of the user lexicon not based on new words or new pronunciation (see col. 6, lines 45-50) (e.g. Since the updating of the language models is performed, the extraction of the specific word will be retrieved and hence is an alternate form of a word in the alternate list as indicated by the reference (e.g. The example given is "steep" and "step")).

Hon *et al.* (801) does teach the use of a lexicon, which is updated for new words (see col. 9, lines 36-40), where words are added (see below) when determining if the words exist in the user lexicon (see col. 7, lines 66-67 and col.

8, lines 1-3) (e.g. The determination is made of whether the word is in the lexicon if it is unrecognized).

As to claim 7, Nassiff *et al.* teaches a method of learning with an automatic speech recognition system, the method comprising:

detecting a change to dictated text (see col. 5, lines 33-40, based on deletion or typing over the words);

inferring whether the change is a correction, or editing (see col. 5, lines 33-48, correction or editing is determined based on deletion (editing) or typing over the words (correction.); and

wherein inferring whether the change is a correction includes comparing a speech recognition engine score (see col. 6, lines 28-31) of the dictated text and of the changed text (see col. 7, lines 50-62).

if the change is inferred to be a correction, selectively learning from the nature of the correction without additional user interaction (see col. 6, lines 45-50).

wherein selectively learning from the nature of correction includes determining if the corrected word exists in the user lexicon, selectively learning the pronunciation (see col. 6, lines 45-50, the language model is updated when the replacement word is found on the alternative word list.)

However, Nasiff does not specifically teach the selectively increase a probability associated with word.

Hon does teach selectively increase a probability associated with word (see col. 2, lines 30-36, where the language model is adapted to increase the chance the same word is recognized by increasing the unigram probability.)

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the correction of dictating speech of Nassiff *et al.* with the inclusion of increasing a probability as taught by Hon (801). The motivation to have combined the references involves the reduction of errors when spoken words are not found in the lexicon of the recognition engine so as to adapt to unrecognized words in a speech recognition system (see Hon (801) col. 1, lines 33-36 and lines 54-56).

8. Claims 14-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nassiff *et al.* in view of Hon *et al.* (US 5,852,801) as applied to claim 7 above, and further in view of Lewis *et al.* (US 6,138,099).

As to claim 14, Nassiff *et al.* and Hon *et al.* (US 5,852,801) do not teach the forced alignment of the wave based on a context word.

Lewis *et al.* does teach doing a forced alignment (see Figure 2, step 40, comparison of original audio an baseform of replacement text) of a wave (see Figure 2, step 36, wave is the text or audio) based on at least one context word if such a word exists (see Figure 2, step 36, 38, and 40, the input text of a speech session is used for comparison and determination of replacement text is made).



It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the correction of dictated speech presented by Nassiff *et al.* and Hon *et al.* (US 5,852,801) with the inclusion of alignment between two words as taught by Lewis. The motivation to have combined the references involves updating language models during speech misrecognition without user interaction (see Lewis, col. 1, lines 25-31) as would benefit the speech recognition system presented by Nassiff *et al.* to enhance phonetic and pronunciation recognition.

As to claim 15, Nassiff in view of Hon teaches all of the limitations as in claim 14.

Furthermore, Lewis *et al.* teaches wherein determining if the user's pronunciation deviated from existing pronunciations includes identifying in the wave the pronunciation (see step 40 and step 42, where the original and corrected baseform are compared to see if deviation exists).

As to claim 16, Nassiff in view of Hon teaches all of the limitations as in claim 1.

Furthermore, Hon *et al.* (US 5,963,903) teaches wherein building a lattice based upon possible pronunciations of the corrected word and the recognition result. (see col. Figure 2, step 50, baseform of replacement text generated if not exists and compares in step 40.) (e.g. Hence, it is obvious that the original audio/text also has a baseform representation in order for comparing the two alignments)

As to claims 17-19, Nassiff in view of Hon teaches all of the limitations as in claim 1.

Furthermore, Lewis *et al.* teaches wherein generating a confidence score based at least in part upon the distance of the newly identified pronunciation the possible pronunciations (see Figure 2, steps 40 and 42, where the two baseforms of the original and replaced text are compared to determine whether an acoustic match occurs. As to claim 18, a close acoustic match between the two texts are determined based on some type of scoring. As to claim 19, in order to propagate from step 42 to 41 or 42, a threshold is needed).

9. Claims 22 and 23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nassiff *et al.* in view of Hon *et al.* (801) as applied to claim 22 above, and further in view of Hoffman *et al.* (US 2003/0139922).

As to claims 22 and 23, Nassiff *et al.* in view of Hon *et al.* teach all of the limitations as in claim 22, above.

Furthermore, Nassiff *et al.* teaches the recognition of two similarly recognizable words (word pair). Specifically, "steep" and "step" as in col. 7, lines 43-60. (e.g. Determination is made if the word is in the replacement word in on the list. If it is not then a close match is found. Each word on the replacement list represents a corresponding pair to another word that may be misrecognized.

Furthermore, the Hon (801) reference was used to teach the adding of a word to a lexicon (see col. 9, lines 36-40).

However, Nassiff in view of Hon *et al.* do not specifically teach addition of a word pair temporarily based on the most recent time the word pair is observed and the relative frequency that the pair has been observed in the past.

Hoffmann *et al.* teaches the addition of a word to a lexicon (vocabulary) is based at least partially upon the most recent time the word pair is observed (see [0015], FIFO, where the words not used for a long time are omitted) and the relative frequency (see [0015] and [0031], frequency of occurrence, that the pair has been observed in the past.)

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech recognition system as taught by Nassiff *et al.* in view of Hon *et al.* with the updating a vocabulary depending on frequency and time as taught by Hoffmann *et al.*. The motivation to have combined the references involves continuous renewal of the vocabulary to eliminate word snot used often and those not used for a long time (See Hoffmann *et al.*, [0015]).

### **Conclusion**

10. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP

§ 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

11. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

LaRue (US 5,748,840) is cited to disclose improving reliability of recognizing words in a large database. Chen et al. (US 5,864,805) is cited to disclose error correction in a dictation system. Waibel et al. (US 5,855,000) is cited to disclose correction of transcribed input using a secondary input. Wright (US 6,195,635) is cited to disclose user-cued speech recognition for improving recognition of repeated utterances.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PARAS SHAH whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:00a.m.-4:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571)272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Paras Shah/  
Examiner, Art Unit 2626

11/14/2008

/Patrick N. Edouard/  
Supervisory Patent Examiner, Art Unit 2626